DOCUMENT RESUME

ED 307 327 TM 013 485

AUTHOR Tamir, Pinchas
TITLE Some Issues Related to the Use of Justifications to Multiple Choice Answers.
PUB DATE 89
NOTE 31p.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Biology; College Entrance Examinations; Comparative Analysis; Cross Cultural Studies; *Distractors (Tests); Foreign Countries; Grade 10; Grade 11; Grade 12; High Schools; *High School Students; *Multiple Choice Tests; *Science Tests; Secondary Education; Test Construction; *Test Format
IDENTIFIERS *Answer Constructs; California; Israel; Justification (Psychology); Parallel Test Forms

ABSTRACT

An investigation of biology matriculation tests in Israel was undertaken to assess the use of justifications with multiple-choice items and to compare the effect of three item formats on students' performance. More specifically, the study was designed to determine the: (1) extent to which justifications will differ if the correct answer is made known to the students so that they can concentrate on the justification; and (2) necessity of distractors. Subject samples included 114 tenth and eleventh graders in 7 California high schools in 1987 and 350 Israeli twelfth graders who studied biology toward the matriculation examination in 1987. The same 20 items were presented to the students in three formats--standard multiple-choice, multiple-choice in which the correct answer is marked, and a simple statement of the correct answer without the distractors. Students' justifications for the three item formats were compared, and significant differences were found. Results confirm the usefulness of justifications as a diagnostic tool and indicate the usefulness of future uses of justifications, particularly in the development of two-tier items. (TJH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED307327

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

PINCHAS TAMIR

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

SOME ISSUES RELATED TO THE USE OF JUSTIFICATIONS TO MULTIPLE CHOICE ANSWERS

Pinchas Tamir

School of Education and Israel Science Teaching Center

Hebrew University Jerusalem

TM013485

BEST COPY AVAILABLE

ABSTRACT

In recent years there has been a growing interest in misconceptions which are prevalent among students and in ways and means which can be used to identify these misconceptions. The present study focuses on the use of justifications to multiple choice items. This approach has been employed for several years in the biology matriculation tests in Israel and found to be a very effective diagnostic tool.

Two specific questions were studied:

1. To what extent will the justifications differ if the correct answer is made known to the students so that they can concentrate on the justifications without being bothered by the need to search for the best answer.
2. Since in the case described in 1 above the correct answer is made known perhaps the distractors are altogether unnecessary?

Two samples were studied. The first comprised of 114 10th and 11th grade students who studied biology in the year 1987 in 7 high schools in California. The second consisted of 350 Israeli 12th grade students who studied biology toward the matriculation examination in the same year.

The same 20 items were presented to the students in three formats, namely: regular multiple choice, multiple

choice in which the correct answer is marked and a statement of the correct answer without the distractors.

Students' justification for the three item formats were compared and significant differences were found.

The study confirmed the usefulness of justifications as a diagnostic tool and offered recommendations regarding future uses of justifications, including their contribution to the construction of two tier items.

Introduction

In recent years there has been a growing interest in students' perceptions and misconceptions. It is widely acknowledged that misconceptions: (a) are quite widespread among students in many subject areas of science; (b) are very persistent and difficult to change or replace; and (c) adversely affect subsequent learning (e.g. Smith and Lott, 1983; Nussbaum and Novick, 1982).

Various means and methods are described in the literature for identifying misconceptions. By far the most widely used is the clinical interview (Gilbert, Osborne and Fensham, 1982; Nussbaum, 1979). This method has the advantage of providing excellent in-depth information about the students' conceptions. However, it has some serious drawbacks: (a) it is time and labor intensive and thus difficult to apply to large numbers of students; (b) the basis for generalizing the findings to large groups of students is rather limited; (c) it cannot be employed by teachers as part of their regular instruction.

Attempts have been made to use paper and pencil tests in order to collect data from more students than can be reached through clinical interviews. Two types of question formats have been used in such studies: open-ended short essay and multiple choice. Combinations of the two formats were used by Brumby (1979). In her study, the same item was presented to the students twice: first as an open-ended and then as a multiple choice question. Differences between the responses to the two

formats were not reported. Wandersee (1983) added a requirement for justifications to some of the items on his Photosynthesis Concept Test.

Staver (1986) also used combination items; completion + multiple choice + justification and completion + essay. He found that achievement was affected by the item format only on unfamiliar or complicated tasks. Haslam and Treagust (1987) and Peterson et al. (1986) followed Tobin and Capie (1981) and designed two-tier items for the explicit purpose of identifying misconceptions. The first tier consists of a multiple choice item and the second contains four possible reasons for choice of the first tier. The student has a double task: first she or he has to choose the correct answer among the options of the first tier and then to follow by choosing a justification to support the first choice.

For some years now Israeli students who take the biology matriculation (Bagrut) exam at the end of year 12 have been asked to supply justifications to the choices made on selected multiple choice items. The items selected for this task are of high cognitive level (comprehension and above).

The justification has been given by the students in the form of short answers usually 2-3 sentences long. The results for the year 1985 (N=2405) indicated that the mean score on the justifications task was about 17 percentage points (more than a full standard deviation) lower than the mean multiple choice

score (Amir et al 1987). Similar results were reported by Yarroch (1983). Based on our experience it may be concluded that the contribution of justifications can be twofold:

1. Identifying misconceptions, missing links and teleological reasoning among students who choose the correct option.
2. Gaining better understanding of the notions and misconceptions held by students who choose the distractors.

Following several years of using justifications to multiple choice items in the matriculation examinations, the following two questions were raised:

1. To what extent will the justifications differ if the correct answer is made known to the students, so that they can concentrate on the justification without being bothered by the search for the correct answer.
2. Since in the case described in 1 above the correct answer is made known, perhaps the distractors are altogether unnecessary?

These questions are the focus of the present study.

To our best knowledge the effect of these kinds of changes in the item format on students' performance as well as on the nature and quality of the justifications provided by the students has not as yet been studied.

Purpose

The purpose of the study is to compare the effect of three item formats on students' performance. The three item formats

are: a regular multiple choice (designated as "regular"); a multiple choice item in which the best answer is marked (designated as "marked"); and a statement of the best answer without distractors (designated as "statement").

The following measures were the dependent variables:

a) Percentage of correct justifications for each of the three formats.

b) The nature of justifications provided for each of the item formats. In addition the correlation between performance on regular multiple choice items and the score on justifications for those items was also studied.

Method

Samples

Two similar studies were carried out. The first study involved 350 12th grade Israeli students who were studying toward the biology matriculation examination in the year 1987. The second study involved 200 10th and 11th grade students who studied biology in the same year in 7 schools in California.

Tests and administration

Six test versions, each constructed from the same 20 biology items selected from Israeli matriculation examinations of previous years, were prepared. Versions 1 and 2 comprised 8 regular multiple choice items and another 7 multiple choice items in which the correct answer was marked. Version 3 and 4 comprised 8

regular multiple choice items and 7 correct statements. Versions 5 and 6 comprised 8 multiple choice items in which the correct answers were marked and 7 correct statements. The following biological topics were covered: energy relationships, regulation, ecological relationships, plants and water, photosynthesis, cell water relationships and genetics.

Each class was divided into three groups, each of which responded to one of the versions. The test was administered by the biology teachers during a regular biology period after explaining to the students the purpose of the study and pointing out that responding to the test constitutes a good preparation for the end of year examinations. The design permitted comparison of performance of the same students on different item formats without spending too much testing time.

Analysis

Responses to 200 Israeli tests, randomly selected were carefully analyzed by two independent evaluators for the purpose of discovering different kinds of justifications, both correct and incorrect. On the basis of this analysis different content and reasoning categories were defined, so that each given justification could be described in terms of the category system designed for a given item. The number of categories for a given item ranged between 4 and 12.

These categories were used in the analysis of all the tests. In this way the frequencies of different notions and conceptions

were obtained. Moreover, it was possible to reconstruct common patterns of reasoning associated with the 20 biological concepts which were included in the study. In addition, each student justification was graded as correct (2 points), partially correct (1 point), or incorrect (0 points).

Results

General performance

Table 1 presents the percentages of correct responses of the two samples.

Insert Table 1 about here

It may be seen that the achievement of the U.S. sample is much lower than that of the Israeli sample. To a large extent this may be due to the early administration time (November) so that the U.S. students, who study their biology just for one year, have not had an opportunity to learn many of the topics included in the tests. The difference between the percentage of correct multiple choice and the percentage of correct justification scores is much smaller in the U.S. sample. Perhaps this may be interpreted as an indication that partial knowledge contributes to the success in multiple choice items more than to success in supplying justifications. This is so because there is a higher probability of getting a correct answer in a multiple choice item by an educated guess whereas a correct justification requires full knowledge.

It may also be seen that the percentage of correct answers varies greatly even within the same topics. The variation is greater among the U.S. sample where ecological relationships stands out both in correct choices and in correct justifications, while cell water relations appears to be most difficult on both measures.

The effect of item format

Table 2 presents the results of the assessment of the justifications on the scale of 0=incorrect, 1=partially correct, and 2=correct.

Insert Table 2 about here

Here we see some interesting trends. Generally, the two samples are quite similar. The lower mean score found in both samples for justifications to regular items compared with marked items could have been expected. However, with regard to statements one might predict even a higher score than in marked items, since the distractors in the marked items can be confusing. Yet, the results indicate that for both samples providing correct justifications to statements is more difficult than providing correct justifications for marked items. However, while the standard score different between the two formats for the Israeli sample is only 1/4 of a standard deviation, for the U.S. sample it amounts to 2/3 of a standard deviation. Perhaps the dominance of multiple choice testing in the U.S. has created special difficulty for students when confronted with such an unfamiliar

task. Several students, indeed, responded "it is correct if you say so." In any event, for both samples a marked multiple choice item appears to be easier to justify than a statement. It may be inferred that either the distractors offer some clues which help the students in supplying reasons for the correct explanation, or that some students do not understand what exactly they have to do under these circumstances since the statement is correct.

Table 3 compares the justification scores by item format.

Insert Table 3 about here

It may be seen that in certain topics (e.g. ecological relationships and cell water relations) there are only very small and inconsistent differences. In other topics such as energy transformation or photosynthesis, the mean score of justifications to regular items is substantially lower than that of the remaining two formats.

Relationships

Table 4 presents intercorrelations for the Israeli sample.

Insert Table 4 about here

The data indicate that whereas in the performance on multiple choice items there are no gender differences, boys perform on justifications better than girls. Justification scores are positively correlated with performance on multiple choice in all test formats.

The lower correlation which appears for marked items could have been expected. On the other hand, it is somewhat surprising that the highest correlation is found between performance on multiple choice and supplying justifications to statements. The relatively high correlations among justification scores in the three item formats indicates that close to 50% of the total variance is common. Because of the small number of subjects in the U.S. sample it was not feasible to calculate correlations by item formats. However, the rank order correlation between performance on the multiple choice items and justification scores for all three item formats together was found to be r=0.47.

Incorrect justifications

Incorrect justifications have been classified into three major categories as follows: (1) Tautology, (2) Teleology and/or anthropomorphism, (3) Content specific misconceptions.

Tautology is defined as a repetition of an idea in other words without imparting additional force or clearness. For example, in item 10, a student justified his choice of one of the multiple choice options, namely, the sources of the corn field crop is CO_2 from the air and water and minerals from the soil by saying "the source of this gain in dry weight are water and minerals from the soil and carbon dioxide from the air."

Teleology is defined as the belief that purpose and expected benefit may be the direct cause of a natural phenomenon. For example: "In daylight the stomates open and water is lost by transpiration. The plant takes in water in order to fill in the lost water" or "the plant gives off more oxygen than CO_2 because otherwise living organisms would not be able to survive"; "glycogen is stored in the liver in order to keep constant the sugar concentration in the blood." Anthropomorphism is defined as ascribing human forms or attributes to beings or things not human. For example: "A mutation takes place to enable an organism ot adjust to the environment," or the "water molecules move in because they try to achieve balance."

The data in Table 5 show that the tautology is used much more frequently by the U.S. than by the Israeli sample.

Insert Table 5 about here

Out of all incorrect justifications tautology takes the lead reaching, on the average, 56%. This result may reflect the situation of U.S. students who have hardly been asked, perhaps, in all their years of schooling, to justify their choices. Perhaps the hidden message has been: "You just have to memorize the correct answer - don't bother beyond that."

Although teleology and anthropomorphism account for only 10% of the wrong justifications in both samples, this figure may be misleading. In fact, if we leave out tautology, they account for 30% of the remaining justifications.

Misconceptions

A major benefit derived from justifications is the opportunity they offer to identify students misconceptions. Following are selected misconceptions which were uncovered in the present study pertaining to four biological topics.

Respiration and photosynthesis

1. In animals the source of energy for their activities is the oxidation of organic compounds while in plants photosynthesis provides the energy for their activities.
2. In the process of respiration organic compounds are broken down to their elements. These elements are the raw material from which ATP is produced.
3. Autotrophic organisms carry out photosynthesis but not respiration.
4. Autotrophic organisms carry out respiration only at night.
5. Plants obtain their food from the soil.
6. Minerals are among the raw materials for photosynthesis.
7. At 4:00 PM it is very hot, hence much water will be lost by transpiration and consequently the dry weight of the leaves will increase.
8. Photosynthesis is the respiration of plants.
9. More oxygen than CO_2 is given off by plants over the 24 hours of a day because there are more daylight than night hours.
10. Plants give off more oxygen than CO_2 because they do not need the oxygen but they do need CO_2 for photosynthesis.

Genetics

1. When a mutation occurs, if it is recessive it will disappear in one or two generations, while a dominant trait will continue to appear in future generations.
2. Survival of a mutant depends on its ability to adapt itself to the environment during its life span.
3. Meiosis is the division of the gametes (rather than of mother cells).
4. The primary structure of an enzyme is determined by the sequence of DNA bases in the enzyme molecule.

Cell water relations

1. The concentration of K^+ ions in an aquatic plant is 40 times as high as in the surrounding water because the cell wall does not allow K^+ ions to move out.
2. Plants need K^+ so they absorb as much as they need.
3. In a plant cell in a state of plasmolysis the space between the cell wall and the protoplast is filled with the solution which surrounds the cell
   a) due to the presence of the outside solution
   b) due to the vacuum which exists there
4. When beet cubes are put in water in 70° C the water turns red because
   a) the beet cells burst in 70° C
   b) the red pigments are destroyed at 70° C
   c) the cell walls are destroyed at 70° C

Regulation

Mammals as tiny as ants do not exit because they cannot maintain a constant body temperature. The reason for this is:

1. there is not enough room for a temperature regulating mechanism.
2. the surface is so small that it does not allow sufficient supply of oxygen and/or other essential materials.
3. there is a need for a critical mass below which it is impossible to maintain a constant temperature.
4. the skin is so thin that it cannot maintain a constant temperature inside the body.

Discussion

Previous studies have shown that justifications to multiple choice items increase significantly the information that test results provide about students' knowledge (e.g. Tamir et al., 1989; Yarroch, 1989). The present study was designed to compare the effect of three item formats on the nature and quality of justifications. The major findings are:

1. More students are able to provide correct justifications when the best answer in a multiple choice item is marked for them. Apparently knowing what the best answer is, cuts out some of the "noise" created by the need to eliminate the distractors.
2. It is easier to provide a correct justification to a marked answer within a multiple choice item than to an identical

statement. Apparently the distractors offer some clues which help in understanding and/or in formulating the justification.

3. U.S. students have a greater difficulty than Israeli students in explaining why a certain statement is correct. This may be attributed to the dominance of multiple choice testing in U.S. science classes.
4. Whereas in the performance on multiple choice items there are no gender differences, boys perform on justifications better than girls. This may suggest that, on the average, boys exhibit more meaningful learning than girls.
5. Performance on multiple choice items has a moderate positive correlation with performance on justifications.
6. The relatively high correlations among justification scores in the three item formats suggests the existence of a substantial common variance in the ability to understand and explain relationships regardless of the item format.
7. Incorrect justifications provide a rich source of students' misconceptions. There are no consistent significant differences in the nature and quality of justifications provided to the three item formats.
8. Whereas Israeli and U.S. students exhibit similar levels of reliance on teleological and anthropomorphic reasoning, the U.S. students are much more inclined to use tautology. This orientation may be another outcome of the dominance of

multiple choice testing in U.S. science classes.

Since tautology as well as teleological and anthropomorphic reasoning represent general, rather than content specific misconceptions they deserve serious attention by teachers. Any treatment of these general misconceptions may be expected to result in substantial improvement in understanding cause effect relationships which would be evident in a better performance on justifications.

9. The list of misconceptions compiled during this study indicates that justifications may serve as an effective diagnostic tool.

There are at least three different ways to use lists of misconceptions such as the one presented in this article. First, the specific misconceptions should be communicated to teachers so that they may deal with them while teaching the relevant topics. Second, the misconceptions can be used as a measure of overestimating students' knowledge. Yarroch (1989) used individual interviews to find out what he calls "item validity," namely, to what extent making a correct choice on a multiple choice item actually reflects knowledge and understanding. Two measures employed in our study can serve this purpose well, while saving the expense and time needed for oral interviews. These measures are the percentage of correct relevant components used in the justifications as well as the

justification scores. The higher the scores on these measures the higher is the item validity. Finally, the specific justifications, whether correct or incorrect, can be used to construct two tier items. This way of utilizing the results of justification analysis requires some elaboration.

According to Treagust (1988) a two tier multiple choice item is structured as follows: "The first part of each item on the test is a multiple choice content question having usually two to three choices. The second part of each question contains a set of four possible reasons for the answer given to the first part. The reasons consist of the correct answer, any identified misconception or misconceptions, together with a simple wrong answer if needed. This second part of each item is developed from students' responses given on an open response question as well as information gathered from the interviews and the literature. When more than one misconception has been given by students earlier, more than one misconception is included in the alternative reason responses" (p. 163).

The advantage of a two tier item over a multiple choice followed by a justification provided by each student is two fold. First, it provides all the benefits of an objective item, such as: quick, easy, if desired even machine scoring; economy of response time; it is not dependent on students writing skills. Second, it enables the inclusion of specific misconceptions and provides detailed information about the frequency distribution of these misconceptions in a given student population.

It may be observed that according to Treagust the first tier of a two tier item contains only two or three choices. Take as an example the following item (Treagust 1988, p. 166, somewhat modified):

Which gas is taken in by green plants in large amounts where there is no light energy at all?

(1) carbon dioxide gas

(2) oxygen gas.

The reason for my answer is because this gas is used in:

(a) photosynthesis which occurs in green plants all the time.

(b) photosynthesis which occurs in green plans when there is no light energy.

(c) respiration which only occurs in green plants when there is no light energy.

(d) respiration which takes place continuously in green plants.

The reason for having usually only two choices in the first tier is that with more choices it would be difficult to provide reasons which would match logically (whether correctly or incorrectly) all the choices.

Based on our study we suggest another option for a two tier item which solves the problem of matching the two tiers. According to this suggestion the first tier will consist of a regular multiple choice choice item which was administered

earlier accompanied by written justifications supplied by the students. Out of these justifications appropriate responses are chosen to form a second tier. When, eventually, used as a two tier item, the best answer in the first tier (the original multiple choice question) is marked, so that the reasons provided by the second tier are well matched to the marked answer.

Following are two examples both dealing with photosynthesis and respiration:

What is the source of energy used by animals to carry out their activities?

1. Digestion of food in the digestive system
2. Taking in oxygen and given out carbon dioxide
*3. Oxidizing organic compounds
4. The activities carried out by the body.

The reason for this being the best answer is:

*1. The energy released during oxidation is used to synthesize ATP.
2. This is true for animals while plants obtain the energy for their activities from photosynthesis.
3. In the process of respiration organic compounds are broken down to their elements. These elements are the raw materials from which ATP is synthesized.
4. In the process of respiration ATP is oxidized and broken down thereby releasing energy.

Fifty discs of a diameter of 1 cm each were cut of a sunflower leaf at 4:00 a.m. (Group A). Another fifty discs, similar to the first ones, were cut on the same day at 4:00 p.m. (Group B) and another fifty discs at 4:00 a.m. the next day (Group C). The discs were dried in an oven at 105° C and weighed. It was found that:

1. Group B had the highest dry weight
2. Group C had the highest dry weight
3. Group B had the lowest dry weight
*4. Group A had the highest dry weight

The reason for this being the best answer is:

1. At 4:00 p.m. it is very hot, hence much water will be lost by transpiration and consequently the dry weight of the leaf will increase.
*2. The leaf carries out photosynthesis during the day so the products accumulate. At night a large proportion of the products is transported to other plant organs.
3. During the day water is utilized as a raw material for the process of photosynthesis and, as a result, the dry weight increases.
4. Since plants carry out respiration only at night, the organic compounds are broken down at night and as a result the dry weight is very low in the morning.

Conclusions and Implications

1. The use of justifications to multiple choice items was found to be a sensitive and effective means of assessing meaningful learning. It is recommended that teachers, examiners and researchers incorporate the two formats described here into their regular assessment repertoire.

2. The use of justifications to simple correct statements is not recommended because its diagnostic potential is substantially lower than that of a good multiple choice item.

3. Regarding the two formats of multiple choice it is suggested that the use of multiple choice items in which the correct answer is marked can be very useful when the purpose is to find out whether students really understand particular phenomena, because some "noise" created by the need to eliminate the distractors is cut out.

4. Justification analysis provides raw material for the construction of effective diagnostic two tier multiple choice items.

5. In addition to the diagnostic value of the justifications they also have a strong "back-wash" effect. Students who know that they may be asked to justify their choices will learn their subject matter in a more meaningful way and in more depth so that they will be prepared to write an adequate and complete justification.

References

Amir, R., Frankl, D. and Tamir, P. (1987). Justifications of answers to multiple choice items as a means for identifying misconceptions. In J.D. Novak (Ed.) Proceedings of the 2nd International Seminar on Misconceptions and Educational Strategies in Science and Mathematics. Cornell University Ithaca, NY, Vol. I, 15-25 .

Brumby, M. (1979). Problems in learning the concept of natural selection. Journal of Biological Education, 13, 119-121.

Gilbert, J.K., Osborne, R.J. and Fensham, P.J. (1982). Children's science and its consequences for teaching, Science Education, 66, 623-633.

Haslam, F. and Treagust, D.F. (1987). Diagnosing secondary students' misconceptions of photosynthesis and respiration using at two tier multiple choice instrument, Journal of Biological Education, 21, 203-211.

Nussbaum, J. (1979). Children's conceptions of earth as a cosmic body: A cross age study. Science Education, 63, 83-93.

Nussbaum, J., Novick, S. (1982). Alternative frameworks, conceptual conflict and accommodation: Toward a principled teaching strategy, Instructional Science, 11 , 183-200.

Peterson, R.F., Treagust, D.F. and Garnett, P. (1986). Identification of secondary students' misconceptions of covalent bonding and structure concepts using a diagnostic test instrument, Research in Science Education, 16, 40-48.

Smith, E.L., Lott, G.W. (1983). Teaching for conceptual change:

Some ways to go wrong. In H. Helm and J. Novak (Eds.) International Seminar on Misconceptions in Science and Mathematics. Cornell University, Ithaca, NY, 57-66 .

Staver, J.R. (1986). The effects of problem format, number of independent variables and their interaction on student performance on a control of variables reasoning problem. Journal of Research in Science Teaching, 23, 533-542.

Tamir, P., Amir, R. and Frankel, D. (1989). Multiple choice questions as a diagnostic tool, Questioning Exchange, in press.

Tobin, K.G. and Capie, W. (1981). The development and validation of a group test of logical thinking, Educational and Psychological Measurement, 11, 413-423.

Treagust, D.F. (1988). Development and use of diagnostic tests to evaluate students' misconceptions in science, International Journal of Science Education, 10, 159-170.

Wandersee, J.H. (1983). Students' misconceptions about photosynthesis: A cross-age study, in H. Helm and J. Novak (Eds.) Proceedings of the International Seminar on misconceptions in Science and Mathematics, Ithaca, NY, Cornell University.

Yarroch, W.L. (1989). The implications of content versus item validity on science tests, Journal of Research in Science Teaching, in press.

Table 1

Percentage of correct responses in multiple choice and justifications

Topic	Item	No. of content	Multiple Choice U.S. N=39	Multiple Choice Israel N=120	Justifications U.S. N=114	Justifications Israel N=374
Energy transformations	1	Energy released through oxidation	21	75	41	62
	2	At rest ATP accumulates in muscles	41	78	55	67
Regulation	3	Glucose concentration in blood	61	81	35	59
	4	Temperature regulation in mammals	27	67	32	57
Ecological relationships	5	Pest-parasites relationship	44	67	57	70
	6	Prey-predator relationship	64	91	63	87
Plants and water	8	Growth in the desert	37	87	68	70
	9	Transpiration rate	23	70	10	65
Photosynthesis	7	Autotrophic organisms	15	68	22	51
	10	Sources of plant crops	58	72	40	52
	11	Rate of photosynthesis during the day	26	68	29	63
	12	CO_2-O_2 exchange in plants	52	74	24	49
Genetics	13	Success of mutations	42	80	38	65
	14	Significance of meiosis	29	83	10	57
	15	Hardy Weinberg law	15	87	6	68
	16	Identifying genotypes by cross hybridization	58	78	25	70
	19	Determination of enzyme structure	46	84	37	59
Cell water relationship	17	Active transport of ions	20	69	17	67
	18	Plasmolysis: reasons for	30	61	15	49
	20	Permeability in red beet	23	80	38	71
		Mean	37	75	31	62

Table 2

Mean justification scores in the three test formats

Test format	U.S. sample N	$\bar{x}$	S.D.	Israeli sample N	$\bar{x}$	S.D.
Regular	111	30.8	15.0	374	59.9	22.0
Marked	114	35.0	15.0	332	64.4	21.7
Statement	111	26.3	12.1	334	64.8	23.7

Table 3

Mean percentage scores of justifications

of the Israeli sample by item format*

Topic	Item No.	Regular N=374	Marked N=332	Statement N=344
Energy transformation	1	58	64	66
	2	58	73	71
	Total	58	69	69
Regulation	3	55	60	62
	4	54	61	54
	Total	55	61	58
Ecological relation-ships	5	68	71	74
	6	88	89	83
	Total	78	80	79
Plant and water	8	67	66	76
	9	63	69	63
	Total	65	68	70
Photosynthesis	7	47	52	54
	10	51	49	58
	11	55	66	69
	12	44	52	49
	Total	49	55	58
Genetics	13	59	67	69
	14	57	57	56
	15	67	61	78
	16	69	74	67
	19	51	63	64
	Total	61	64	67
Cell water relations	17	62	69	72
	18	52	52	44
	20	72	73	68
	Total	62	65	61

*The percentage is based on the total number of justifications supplied by the students

Table 4

Intercorrelations among correct multiple choice, correct justifications in the three test formats and gender

	Multiple choice N=374	Justifications Regular N=374	Marked N=332	Statement N=344
Gender	.01	-.15*	-.17*	-.17*
Correct multiple choice		.33*	.24*	.43*
Justification Regular			.64*	.64*
Justification Marked				.77*

*p 0.0

Table 5
Relative frequency of tautology, teleology and anthropomorphism in the incorrect justifications of the two samples (in percentages)

Topic	Item No.	Tautology Israel N=374	Tautology USA N=374	Teleo. and Anthro. Israel N=374	Teleo. and Anthro. USA N=114
Energy transformation	1	28	56	8	14
	2	32	53	21	7
	Total	30	55	15	11
Regulation	3	22	34	8	10
	4	25	53	3	7
	Total	24	44	6	9
Ecological relationships	5	4	49	3	21
	6	10	30	11	0
	Total	7	40	7	11
Plants and water	8	7	39	48	14
	9	6	64	20	10
	Total	7	52	34	12
Photosynthesis	7	4	60	15	8
	10	36	64	5	10
	11	16	35	56	7
	12	13	45	21	24
	Total	17	51	24	12
Genetics	13	24	62	7	15
	14	14	72	14	12
	15	40	70	0	0
	16	38	73	0	7
	19	46	78	0	0
	Total	32	73	4	7
Cell water relations	17	25	60	13	9
	18	21	64	6	16
	20	33	56	5	9
	Total	26	60	8	11
	Grand mean	22	56	11	10